

---

# AP Statistics

## Practice Examination 3

### *Multiple Choice*

### *Statistics*

### *Section I*

---

**Time: 90 minutes**

1. According to the U.S. Bureau of the Census, the distribution of ages for senior citizens was as follows in 2000:

Age	65–74	75–84	85–94	95 and over
Percentage	58.0	32.2	9.0	0.8

Of a random sample of 626 senior residents from a large community, 300 are ages 65–74, 232 are ages 75–84, 92 are ages 85–94, and 2 are over 95. Is there evidence to show that the distribution of residents in the community is different from the distribution of senior citizens in the nation?

- (A) There is evidence of a significant difference in distributions at the 0.10 level but not at the 0.05 level.
- (B) There is evidence of a significant difference in distributions at the 0.05 level but not at the 0.01 level.
- (C) There is evidence of a significant difference in distributions at the 0.01 level but not at the 0.001 level.
- (D) There is evidence of a significant difference in distributions at the 0.001 level.
- (E) There is insufficient evidence of a significant difference in distributions at any of the commonly accepted levels.

2. A high school administrator wishes to compare state assessment results of male and female students.  
Which graphical display *wouldn't* you recommend to him?
- (A) Comparative dotplots
  - (B) Back-to-back stemplot
  - (C) Scatterplot
  - (D) Parallel boxplots
  - (E) Histograms drawn side by side with the same scale
3. In a small town of 500 the last census reported the average family income as \$34,000. Before the next census the town wants to determine whether its average income has significantly increased. To accomplish this, a town administrator could
- (A) randomly sample 70 families, compute the average income, and perform a one-sample  $z$ -test.
  - (B) randomly sample 50 families, compute the average income, and perform a one-sample  $t$ -test with 50 degrees of freedom.
  - (C) randomly sample 50 families, compute the average income, and perform a one-sample  $t$ -test with 500 degrees of freedom.
  - (D) randomly sample 30 families, compute the average income, and perform a one-sample  $t$ -test with 29 degrees of freedom.
  - (E) randomly sample 30 families, compute the average income, and perform a one-sample  $t$ -test with 499 degrees of freedom.
4. Two random variables,  $X$  and  $Y$ , are independent.  $X$  has expected value 2.5 and standard deviation 0.3 while  $Y$  has expected value 4.7 and standard deviation 0.4. Which of the following is true?
- (A) The mean of  $X + Y$  is 6.2.
  - (B) The standard deviation of  $X + Y$  is 0.7.
  - (C) The variance of  $X + Y$  is 0.7.
  - (D) The mean of  $X + Y$  is 11.75.
  - (E) The standard deviation of  $X + Y$  is 0.5.
5. Which of the following is true?
- (A) The value of a random variable must always be positive.
  - (B) The expected value of a random variable must always be positive.
  - (C) The variance of a random variable must always be positive.
  - (D) The expected value of a random variable must be nonzero.
  - (E) The variance of a random variable must be nonnegative (0 or positive).

6. In 2000, women comprised 15.1% of the Army's active duty troops. In a random sample of 100 recent graduates who enlisted in the Army, 20 were women. Construct a 95% confidence interval for the proportion of women who currently enlist in the Army, and determine whether or not this proportion is significantly different from the proportion in 2000 at the  $\alpha = 0.05$  level.

(A)  $0.151 \pm 1.96\sqrt{\frac{(0.151)(0.849)}{100}};$

since 0.151 is contained in the 95% confidence interval, there is insufficient evidence to show that this proportion is significantly different than in 2000.

(B)  $0.20 \pm 1.645\sqrt{\frac{(0.20)(0.80)}{100}};$

since 0.151 is contained in the 95% confidence interval, there is insufficient evidence to show that this proportion is significantly different than in 2000.

(C)  $0.20 \pm 1.645\sqrt{\frac{(0.20)(0.80)}{100}};$

since 0.151 is contained in the 95% confidence interval, there is sufficient evidence to show that this proportion is significantly different than in 2000.

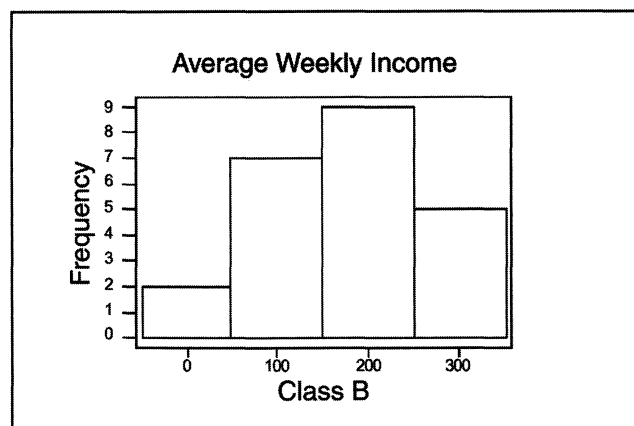
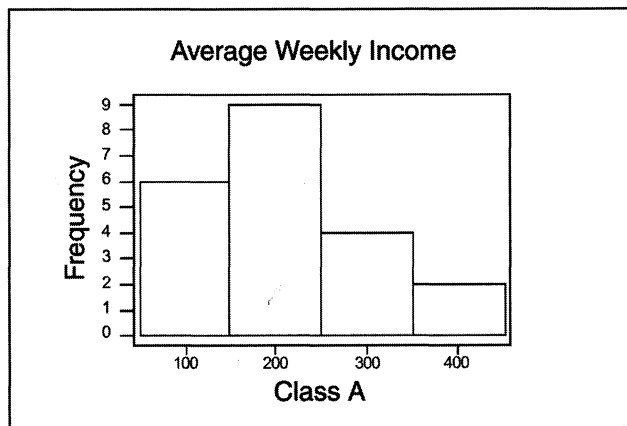
(D)  $0.20 \pm 1.96\sqrt{\frac{(0.20)(0.80)}{100}};$

since 0.151 is contained in the 95% confidence interval, there is insufficient evidence to show that this proportion is significantly different than in 2000.

(E)  $0.20 \pm 1.96\sqrt{\frac{(0.20)(0.80)}{100}};$

since 0.151 is contained in the 95% confidence interval, there is sufficient evidence to show that this proportion is significantly different than in 2000.

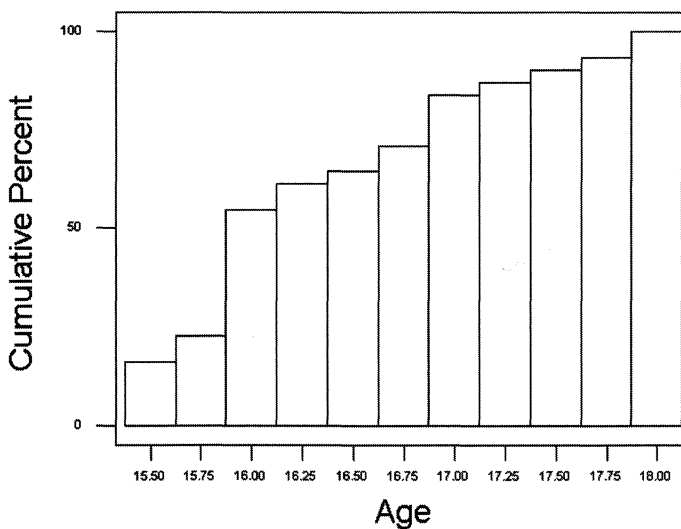
7. Home pregnancy test kits have grown in popularity. Research shows that only 30% of those using a particular kit are actually pregnant. When a pregnant woman uses this kit, it correctly indicates pregnancy 96% of the time. A woman who is not pregnant gets a correct indication 90% of the time. What is the probability that a woman is pregnant given that this test gives a positive result?
- (A) About 96%
- (B) About 86%
- (C) About 80%
- (D) About 36%
- (E) About 21%
8. Which of the following statements is true for the construction of a bar graph?
- (A) The scale of the horizontal axis should be the frequency.
- (B) There can be no gaps between bars.
- (C) Bar graphs are used to depict categorical data.
- (D) Mention of the shape, center, and spread should be made when describing bar graphs.
- (E) The width of the bars should differ according to the frequency of the class.
9. The histograms below represent average weekly job income for students in two high school classes.



What conclusion is reasonable based on the displays? *Note:* Both classes have medians in the \$200 range.

- (A) The mean weekly income for class A is higher than for class B.
- (B) The mean weekly income for class B is higher than for class A.
- (C) More students in class B earn wages in excess of \$250 than in class A.
- (D) More students in class B have higher paying jobs than in class A.
- (E) All students in both classes have paying jobs.

10. Two friends, Tom and Janice, have cars in desperate need of repair. On any given day, the probability that Tom's car will break down is 0.5, the probability that Janice's car will break down is 0.5, and the probability that both of their cars will break down is 0.3. What is the probability that Tom or Janice's car will break down?
- (A) 1.3  
(B) 1.0  
(C) 0.7  
(D) 0.4  
(E) 0.2
11. In a state where young adults can obtain learner's permits for driving 6 months before their 16th birthdays a random sample of young adults applying for their permits is taken, and their ages in years are recorded. The cumulative proportions are plotted against age, resulting in the following graph.



The median and interquartile range for this group of individuals are

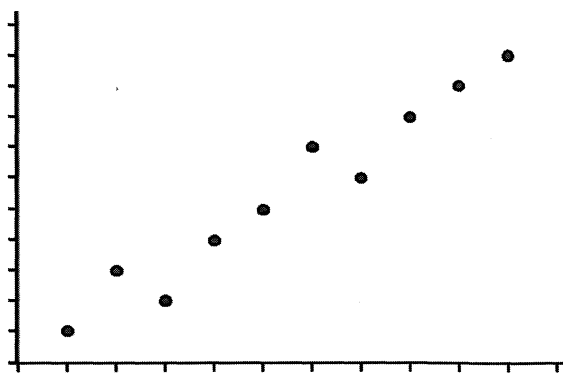
- (A) 16 and 1 year.  
(B) 16 and 1.5 years.  
(C) 16.75 and 1 year.  
(D) 16.75 and 1.5 years.  
(E) 50 and 50 years.

12. A local senior center would like to estimate the average life expectancy for men in the local community accurate to within 0.5 year. Assuming a standard deviation of 10 years and a confidence level of 90%, how many local male death records would need to be sampled?
- (A) 33
  - (B) 658
  - (C) 1083
  - (D) 1536
  - (E) 108,241
13. In general, how does quadrupling the sample size affect the width of a confidence interval?
- (A) The width of the interval becomes four times as large.
  - (B) The width of the interval becomes two times as large.
  - (C) The width of the interval becomes half as large.
  - (D) The width of the interval becomes one-quarter as large.
  - (E) We need to know the sample size to be able to determine the effect.
14. The  $p$ -value for a significance test is 0.0358. A correct interpretation of this  $p$ -value would be:
- (A) About 3.6% of the data are significant.
  - (B) About 3.6% of all samples are significant.
  - (C) About 3.6% of all samples would produce a test statistic at least as extreme as ours if the null hypothesis is true.
  - (D) There is sufficient evidence to reject the null hypothesis.
  - (E) There is sufficient evidence to fail to reject the null hypothesis.

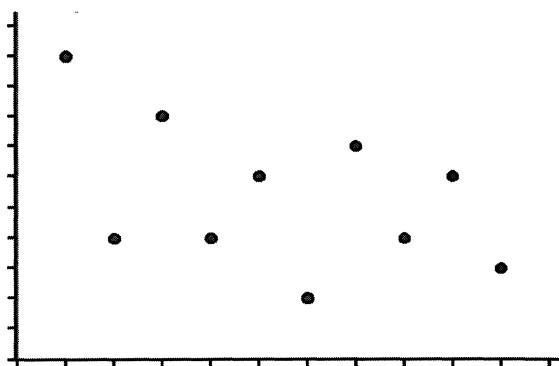
15. In order to ease parking problems in a community containing a university, university officials propose purchasing one acre of community parkland that is adjacent to the university to build a parking garage. The officials believe community members will overwhelmingly support this proposal, and they would like to conduct a survey of 100 community members to confirm their belief. Which of the following will produce a simple random sample?
- (A) Recording the opinion of the first 100 people who call the university regarding this issue
  - (B) Randomly selecting 100 people from the local phone directory
  - (C) Surveying every third person who walks past the administrative offices until 100 people have responded
  - (D) Using the latest census data from the community, numbering the residents, and using a random-number table to choose 100 people
  - (E) Using the latest census data from the community and randomly choosing 25 residents ages 18–25, 25 residents ages 26–39, 25 residents ages 40–64, and 25 residents ages 65 and over
16. Summary statistics are calculated for a data set that includes an outlier. If the outlier is removed, which summary statistic would be least affected?
- (A) Mean
  - (B) Median
  - (C) Range
  - (D) Standard deviation
  - (E) Variance
17. Recently, news channels in Montana and California randomly polled adults in their states about their belief that air quality was negatively impacting their health. The results showed that 5% of the adults in Montana believed that air quality was a serious problem, while in California the results showed that 43% believed the issue was serious. The reporter concluded that 24% of all adults from these two states believe there is a problem with air quality. Which of the following statements best describes the problem with this reporter's conclusion?
- (A) We do not know if a simple random sample was conducted in Montana and California.
  - (B) The reporter should have stated that 48% of all adults believe there is a problem with air quality.
  - (C) The size of the population in Montana differs from that in California.
  - (D) The reporter should have included a statement about the margin of error with the result.
  - (E) There is no problem with the reporter's conclusion.

EXAM 3

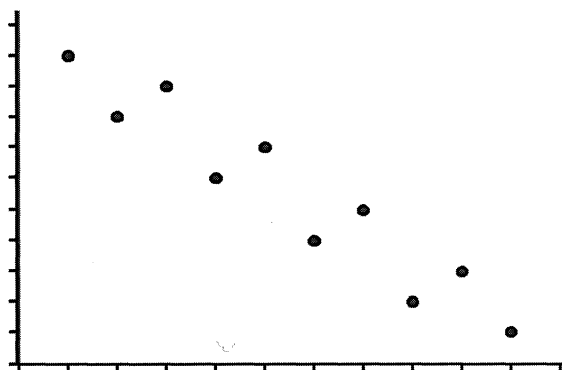
18. Order the correlation coefficients from least to greatest for the given scatterplots.



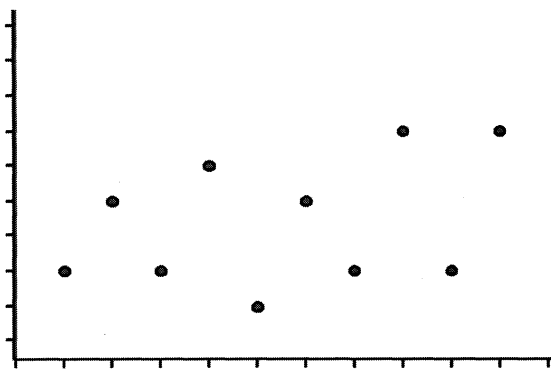
Plot 1 with correlation  $r_1$



Plot 2 with correlation  $r_2$



Plot 3 with correlation  $r_3$

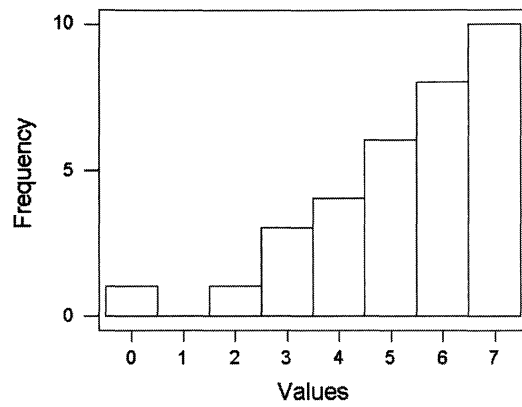


Plot 4 with correlation  $r_4$

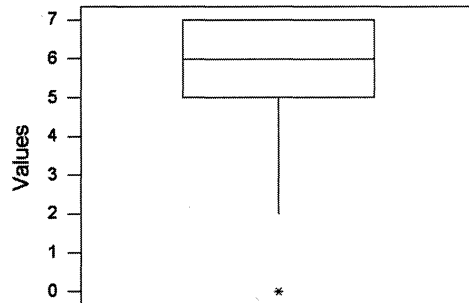
- (A)  $r_4 < r_3 < r_2 < r_1$
- (B)  $r_4 < r_2 < r_3 < r_1$
- (C)  $r_3 < r_2 < r_4 < r_1$
- (D)  $r_2 < r_3 < r_4 < r_1$
- (E)  $r_1 < r_2 < r_3 < r_4$



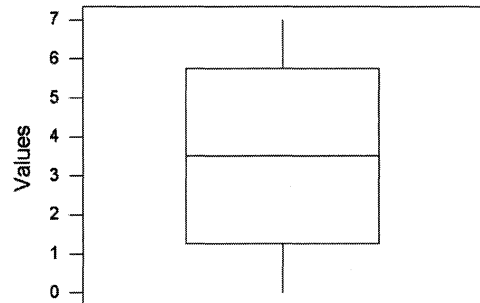
19. Which boxplot matches the given frequency histogram?



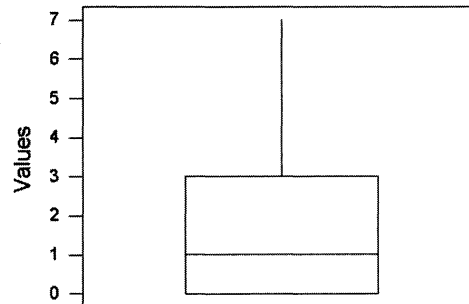
(A)



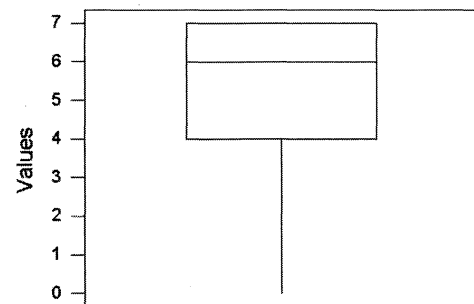
(B)



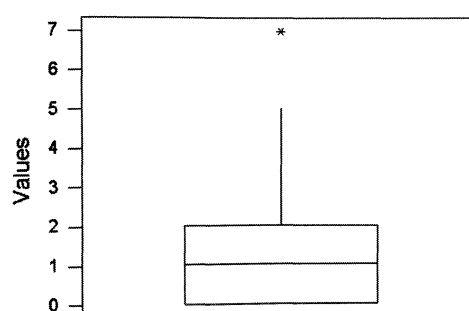
(C)



(D)



(E)



20. Which of the following statements is true?

- (A) A census is an experiment that involves the entire population.
- (B) A parameter is a value used to describe a sample.
- (C) A sample is the entire group of individuals we want information about.
- (D) In stratified random sampling, every individual has the same probability of being chosen.
- (E) Voluntary samples never introduce bias.

21. According to the U.S. Bureau of the Census, in 2000 the breakdown of the U.S. population (in %) by age was as follows.

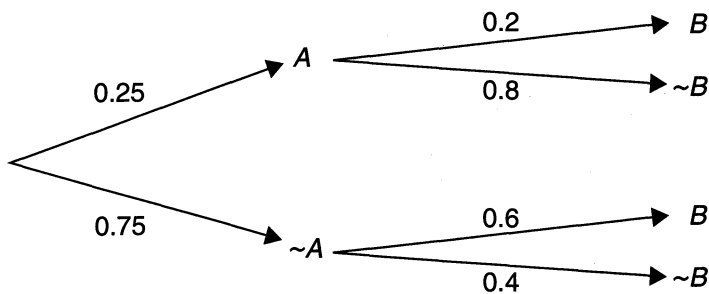
Age	< 5	5–14	15–24	25–34	35–44	45–54	55–64	65–74	75–84	> 85
Percent	6.8	14.6	13.9	14.2	16	13.4	8.6	6.5	4.4	1.5

In order to determine if the distribution of the ages in a city are comparable to those of the country, a random sample of 200 ages are found. What test should be conducted?

- (A) One-sample  $t$ -test
  - (B) Two-sample  $t$ -test
  - (C) Matched-pairs  $t$ -test
  - (D)  $\chi^2$  goodness-of-fit test
  - (E) Inference for regression  $t$ -test
22. A study was conducted to determine the benefit of an over-the-counter medication in reducing the development of disease. Subjects selected were chosen because they were known to be in a high-risk group for the disease. The results of the study are
- (A) not replicable.
  - (B) applicable only to the subjects in the study.
  - (C) not readily generalizable.
  - (D) false and misleading.
  - (E) valid for all takers of this over-the-counter medication.
23. The power of a significance test against a particular alternative is 82%. Which of the following is true?
- (A) The probability of a Type I error is 82%.
  - (B) The probability of a Type I error is 18%.
  - (C) The probability of a Type II error is 82%.
  - (D) The probability of a Type II error is 18%.
  - (E) The probability of making either a Type I error or a Type II error is 18%.

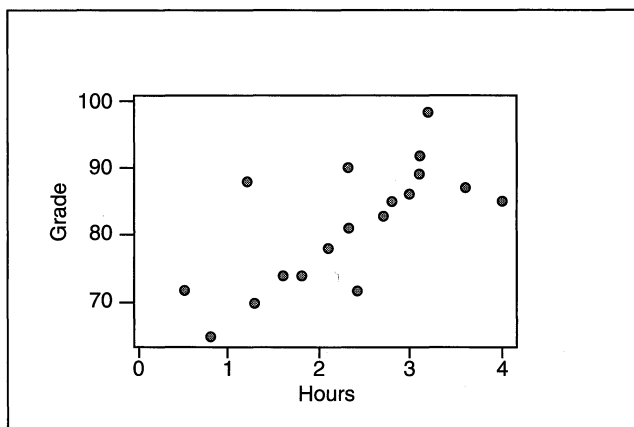
24. A  $t$ -distribution with 30 degrees of freedom is an appropriate statistical model when
- (A) constructing a confidence interval based on a random sample of size 29.
  - (B) constructing a confidence interval based on two independent random samples of sizes 13 and 17.
  - (C) using a  $t$ -statistic based on a random sample of size 30.
  - (D) using a  $t$ -statistic based on a random sample of size 31.
  - (E) we do not know  $\mu$ ,  $\sigma$ , or the sample size.
25. A random sample of 32 games is chosen for a professional basketball team, team A, and their results are recorded. The team averaged 88 points per game with a standard deviation of 8. The same is done for a second team, team B, with this team averaging 90 points per game with a standard deviation of 6. A 95% confidence interval is constructed for the difference in points scored per game between the two teams. What do the results of the confidence interval show?
- (A) We can be 95% confident that on average, team A scores between 1.54 and 5.54 more points per game than team B.
  - (B) We can be 95% confident that on average, team A scores between 1.54 and 5.54 fewer points per game than team B.
  - (C) We can be 95% confident that on average, team A scores between 1.54 points fewer than and 5.54 points more than team B.
  - (D) We can be 95% confident that on average, team A scores between 5.54 points fewer than and 1.54 points more than team B.
  - (E) The conditions necessary to find a 95% confidence interval have not been met.
26. What is the probability that on five rolls of a fair die you will roll three or more 1s?
- (A) 99.7%
  - (B) 96.5%
  - (C) 40.2%
  - (D) 3.5%
  - (E) 0.3%

27. A class of statistics students scored the following grades on a 10-point quiz: 5, 2, 4, 6, 7, 9, 8, 10, 10, 6, 7, 7, 10, 4, 2. Their teacher, who sets a passing grade at 70%, looks at the distribution and sees that it is
- (A) skewed right with nearly 50% of the class passing.
  - (B) skewed left with nearly 50% of the class passing.
  - (C) skewed right with nearly 50% of the class failing.
  - (D) nearly symmetric with one outlier.
  - (E) nearly symmetric with no outliers.
28. A 95% confidence interval for the proportion of female athletes in college programs is constructed based on sample data from 30 randomly selected coeducational colleges. If the confidence interval is (0.38, 0.52), we can say
- (A) 95% of the time, colleges have between 38% and 52% female athletes.
  - (B) 95% of colleges have an average of 45% female athletes.
  - (C) 95% of the time, the true proportion of female college athletes is somewhere between 38% and 52%.
  - (D) we are 95% confident that all colleges have between 38% and 52% female athletes.
  - (E) we are 95% confident that the proportion of female athletes in college programs is between 38% and 52%.
29. Given the tree diagram shown, what is the probability that event B will happen given that event A has already occurred?



- (A) 10%
- (B) 20%
- (C) 45%
- (D) 50%
- (E) 75%

30. Patients afflicted with a debilitating disease took part in a study to measure the effectiveness of a new drug in controlling the progress of the disease. The patients were divided into two groups: an experimental group who received the drug, and a control group who received a placebo. The results of the experimental group were so positive that the study was stopped early. This was most likely because
- (A) the researchers stopped getting useful information.
  - (B) the researchers realized that their subjects were poorly chosen.
  - (C) the researchers felt that it was too expensive to continue the study.
  - (D) the researchers felt it was unethical to use only patients who had the disease in the study.
  - (E) the researchers felt it was unethical to withhold an effective treatment from the placebo group.
31. Statistics students decided to analyze the relationship between hours spent studying per night and average grades on tests. Their data display is shown below.



The correlation coefficient for the least squares regression line is most nearly

- (A) 0.95
- (B) 0.70
- (C) -0.55
- (D) -0.30
- (E) 0.10

32. The weekly beef consumption for a random sample of 15 adults is 1.25 lb with a standard deviation of 0.39 lb. A modified boxplot of the data reveals a slight skew with no outliers. Find a 98% confidence interval for the average weekly beef consumption of adults.
- (A)  $0.39 \pm 2.326 \frac{1.25}{\sqrt{15}}$
- (B)  $0.39 \pm 2.624 \frac{1.25}{\sqrt{15}}$
- (C)  $1.25 \pm 2.326 \frac{0.39}{\sqrt{15}}$
- (D)  $1.25 \pm 2.602 \frac{0.39}{\sqrt{15}}$
- (E)  $1.25 \pm 2.624 \frac{0.39}{\sqrt{15}}$
33. The president of a borough council is trying to determine whether or not there is a need to build a parking garage rather than a parking lot on borough property. The president conducts a survey of borough residents to assess their parking needs. What is the population of interest?
- (A) All state residents
- (B) All borough residents who own cars
- (C) All borough residents who need additional parking
- (D) All visitors to the borough
- (E) Residents of the borough and nearby communities
34. A pollster working on an issue of national importance wants to be sure that the percentage of people with a certain opinion differs by no more than 3%. What sample size should he use for the poll?
- (A) 9
- (B) 17
- (C) 278
- (D) 556
- (E) There is not enough information to determine sample size.

35. A study randomly assigned patients to treatment groups to determine the effect of taking aspirin in preventing the development of colon polyps. One group took an aspirin daily and the other group took a placebo. Neither the patients nor the doctors knew who was getting which pill. This study is best described as a
- (A) block design with random assignment.
  - (B) double-blind comparative experiment.
  - (C) blinded block design observational study.
  - (D) blind experiment with random assignment.
  - (E) randomly assigned observational study.
36. There are 100 students enrolled in various AP\* courses at Addison High School. There are 31 students enrolled in AP\* Statistics, 52 students enrolled in AP\* English, and 15 students enrolled in AP\* French. Ten students study both AP\* Statistics and AP\* English, 5 students study both AP\* Statistics and AP\* French, 8 students study both AP\* English and AP\* French, and 3 students study all 3 AP\* subjects. What is the probability that a student takes an AP\* course other than these 3?
- (A) 9%
  - (B) 12%
  - (C) 22%
  - (D) 78%
  - (E) 93%
37. Two professors, A and B, got into an argument about who grades tougher. Professor A insisted that his grades were lower than those for Professor B. In order to test this theory, each professor took a random sample of 25 student grades and conducted a test of significance. The graphical displays showed that each grade distribution was approximately normal. The results are recorded below.

$H_0$ : Population mean of **Professor A** equals that of **Professor B**  
 $H_a$ : Population mean of **Professor A** is less than that of **Professor B**

	<b>Professor A</b>	<b>Professor B</b>
Count:	25	25
Mean:	79	82
Std dev:	6	4
Std error:	1.2	0.8

Using unpooled variances  
 Student's  $t$ : -2.08  
 DF: 41.8144  
 P-values: 0.022

Which of the following conclusions is/are supported by the results of the significance test?

- I. At the  $\alpha = 0.05$  level, we have evidence to show that every student in Professor A's class scored lower than every student in Professor B's class.
- II. If there were no difference in grades between the two professors, then we could get results as extreme as those from the samples approximately 2.2% of the time.
- III. The test results are not valid, since the conditions necessary to perform the test were not met.

- (A) I only
- (B) II only
- (C) III only
- (D) I and II
- (E) II and III

38. An athlete has recorded his heart rate (in beats per minute) during training for several years. He wishes to determine his average training heart rate. In order to do so, he randomly selects heart rates from 15 days and calculates a 95% confidence interval from the data. If the rates are 111, 116, 118, 119, 127, 129, 132, 132, 134, 135, 138, 141, 146, 152, and 182, what is the 95% confidence interval for his average heart rate during training?

- (A)  $134 \pm 1.645 \frac{17.431}{\sqrt{15}}$
- (B)  $134 \pm 1.761 \frac{17.431}{\sqrt{15}}$
- (C)  $134 \pm 1.960 \frac{17.431}{\sqrt{15}}$
- (D)  $134 \pm 2.145 \frac{17.431}{\sqrt{15}}$
- (E) The conditions for finding a 95% confidence interval have not been met.



39. Regression analysis is performed on two variables, X and Y. A residual plot is graphed, and summary statistics are found for the residuals.

Dependent variable is: Y

No Selector

R squared = 83.0% R squared (adjusted) = 80.5%

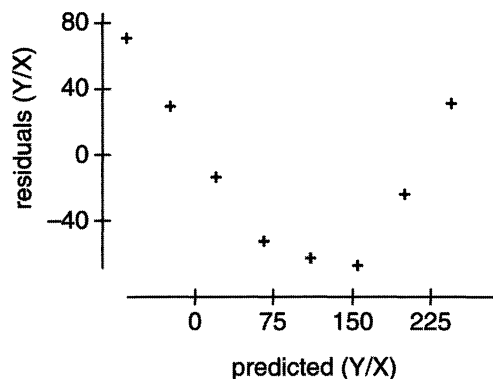
s = 59.67 with 9 - 2 = 7 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	121320	1	121320	34.1
Residual	24921.5	7	3560.21	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	-113.056	43.35	-2.61	0.0350
X	44.9567	7.703	5.84	0.0006

Summary of residuals(Y/X)  
No Selector

Count	9
Mean	-42.6326e-15
Median	-11.8444
MidRange	5.80556
StdDev	55.8139
Range	145.1
IntQRRange	97.3167



What tells you there is a better fit for the data?

- (A) The mean residual value is 0.
- (B) A significant relationship does not exist for X and Y.
- (C) There is a low correlation between X and Y.
- (D) There is a pattern in the residual plot.
- (E) The linear regression model for X and Y is appropriate; a better fit does not exist.

**EXAM 3**

40. The following information resulted from regression analysis for the percentage of white children under the age of 18 versus the percentage of black children under the age of 18 who live in poverty in several large cities.

Dependent variable is: **Black**

No Selector

R squared = 79.5%      R squared (adjusted) = 77.8%

s = 1.072 with 14 - 2 = 12 degrees of freedom

Source	Sum of Squares	df	Mean Square	F-ratio
Regression	53.5882	1	53.5882	46.6
Residual	13.7889	12	1.14907	

Variable	Coefficient	s.e. of Coeff	t-ratio	prob
Constant	30.4954	1.942	15.7	≤ 0.0001
White	0.911941	0.1335	6.83	≤ 0.0001

If the data were reevaluated using the percentage of white children as the dependent variable, the correlation coefficient would be which of the following?

- (A) 0.778
- (B) 0.795
- (C) 0.882
- (D) 0.892
- (E) Not enough information is given to calculate the correlation coefficient.